DOCUMENT RESUME

ED 302 069 FL 017 692

AUTHOR Saleemi, Anjum P.
TITLE Language Learnability and Empirical Plausibility: Null Subjects and Indirect Negative Evidence.
PUB DATE Aug 88
NOTE 8p.; In: Papers and Reports on Child Language Development, Volume 27; see FL 017 572.
PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Child Language; *Difficulty Level; Grammar; *Language Acquisition; *Language Processing; Language Research; Learning Processes; Linguistic Theory
IDENTIFIERS *Learnability Theory

ABSTRACT

Children's ability to learn aspects of their language in the absence of supportive evidence is discussed. Specifically, the learnability of null subjects in languages in which they appear is examined when indirect negative evidence is present. It is concluded that parameters such as the null subject parameter may not generate languages, strictly speaking, but may only fix the bounds within which languages can be realized. It is also suggested that children may be sensitive to the nonoccurrence of certain types of evidence, and that the view of learnability outlined here may be empirically plausible only to the extent that the theory of principles and parameters in language is empirically plausible. (MSE)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

PRCLD 27 (1988)

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
E. Clark
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
This document has been reproduced as received from the person or organization originating it.
Minor changes have been made to improve reproduction quality.
Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

ED302069

LANGUAGE LEARNABILITY AND EMPIRICAL PLAUSIBILITY
Null Subjects and Indirect Negative Evidence

Anjum P. Saleemi
University of Essex

In this paper the linguistic framework assumed is Chomsky's principles-and-parameters theory (e.g. Chomsky 1981, 1986a). First I outline a reanalysis of the null subject parameter; this reanalysis is meant to be illustrative rather than definitive. Then I examine the parameter to determine its learnability, taking Wexler and Manzini's set-theoretical approach (Manzini and Wexler 1987, Wexler and Manzini 1987) as a first approximation to a theory of parameter fixation.

1. Reanalysis

In what follows I assume, with little further comment, that pro-drop and postverbal subjects originate from separate, if not entirely independent, parameters (cf. Hyams 1986, Safir 1985). Obviously, most alleged consequences of pro-drop, e.g. lack of *that*-trace effects and long-distance movement of subjects, can be demonstrated to follow from the inversion of subjects, rather than from pro-drop (cf. Rizzi 1982, but see Picallo 1984 for a contrary view). I further assume, following Chomsky (1986b) and others, that a requirement of proper government (i.e. ECP) does not hold of the pronominal empty category *pro*, which is standardly supposed to occupy the lexically null subject position in pro-drop languages (e.g. Rizzi 1982, Chomsky 1982). Finally, Case, rather than Agr, is considered to be the factor licensing a null subject (Rizzi 1986, Safir 1985); the function of rich inflection is hypothesized to be the identification of the grammatical features of a null subject, a process that may determine whether the latter will be realized in a *particular* language. (See Saleemi, in preparation, for further discussion on these matters.)

Rizzi (1986) argues that licensing and identification of content are characteristically independent processes in the theory of grammar (also see Rizzi 1982). To illustrate, PRO is licensed by the relevant clause of the Generalized Empty Category Principle (Chomsky 1981), but it acquires its content by means of control mechanisms. Likewise, NP-trace and *wh*-trace are sanctioned by the ECP (by lexical government, to be precise), whereas it is through the formation respectively of A-chains and A'-chains that their referential content is fixed. In general, licensing is predominantly syntactic and is exercised through a government-type relation; on the other hand, the recovery of content usually takes place on the LF side, by means of a binding-type relation. In a similar vein, Aoun et al (1987) argue that the ECP, viewed as a disjunctive condition consisting of a lexical government component and a coindexing component, is a spurious generalization

Finally, Italian permits referential pro-drop as well as nonreferential pro-drop, as demonstrated below.

(3) a. (Io) vado al cinema
(I) go to the movies
'I go to the movies'
b. Sembra che Gianni sia molto infelice oggi
seems that John is very unhappy today
'it seems that John is very unhappy today'
c. Piove molto durante il mese di febbraio
rains a lot during the month of February
'it rains a lot during the month of February' (Hyams 1986)

Clearly, a revision of the traditional view of pro-drop is in order. Consequently, in part following some typological observations of Rizzi (1986) and Travis (1984), the parameter is reformulated below as a multi-valued parameter. Before introducing the reformulation of the null subject parameter, the following Case Assignment Principle is adopted (cf. Fabb 1984:43).

(4) *Case Assignment Principle*
An NP must be assigned Case at some level (other than D-structure).

This principle makes it possible for an NP to be Case-marked either at S-structure or at LF. It covers the assignment of Case to null NPs as well as overt NPs, the former acquiring Case only at LF. A key idea, due to Bouchard (1984), is that in pro-drop languages Case assignment can be optionally delayed until LF. Essentially, this idea presupposes that obligatory Case at S-structure requires an NP to be lexically realized at PF, whereas in the event of optionality of syntactic Case an NP need not be so realized, unless some other feature (e.g. focus) forces it to acquire Case in order to be overt. The null subject parameter can now be stated as follows.

(5) *Null Subject Parameter*
The assignment of Case to σ may be delayed until LF; where σ, σ a subject, represents
(a) Ø; or
(b) nonargument; or
(c) nonreferential argument; or
(d) any argument whatsoever.

English and French are associated with value (a) of the parameter, allowing no null subjects. On the other hand, German takes value (b), that permits only nonarguments to be omitted, requiring all argumental subjects to be lexically expressed. Yiddish, Malagasy and Icelandic take value (c), that allows the omission of quasi-

arguments as well as nonarguments, ie all nonreferential subjects. Finally, languages like Italian and Spanish (and possibly also Chinese and Japanese) are associated with value (d), under which any subject, referential or nonreferential, may remain null.

The null subject parameter as stated in (5) gives rise to two related problems, a descriptive problem and a learnability problem, to which I now turn.

2. Descriptive Adequacy

First the descriptive problem. The parameter predicts that pleonastic pro-drop will be optional, just like referential pro-drop. However, whereas referential pro-drop is optional in general, nonreferential pro-drop seems to be mandatory in most pro-drop languages, with a few exceptions, e.g. Irish (Travis 1984:231ff) and Welsh. Practically, then, the pro-drop option with respect to pleonastics might be no more than a Hobson's choice. But there appears to be a simple solution to the problem.

In the spirit of Chomsky's (1981) Avoid Pronoun Principle, the absence of lexical pleonastics in pro-drop languages can be attributed to their pragmatic infelicity. It can be assumed that since pleonastic subjects are nonreferential, they would be superfluous in a pro-drop language. Syntactically, if the pro-drop property indeed springs from the optionality of Case at S-structure, then the avoidance of pleonastic elements is understandable as they will tend not to appear in syntax merely as a spell-out of Case. This will then account for the lack of expletive subjects in most pro-drop languages, and one could still claim (5), pragmatically qualified, to be formally correct. In any event, this maximally general statement of the parameter may be required for languages in which nonreferential pro-drop is in fact optional in many configurations. Nevertheless, due to the irregular distribution of pleonastics in languages, the shortfall in the data available to the learner leaves us with a learnability problem, particularly in the total absence of any kind of negative evidence, as is demonstrated in the following section.

3. Learnability

The problem is that the data the child will get are not exactly the data it will expect under the parameter (5). It is plainly evident that in principle the four values of the parameter should generate languages (i.e. sets of sentences) which form a subset hierarchy, as each value increases the set of well-formed structures allowed by the parameter. This would be compatible with the monotonic model of parameter fixation proposed by Wexler and Manzini (op. cit.), indicating that the parameter could be straightforwardly fixed on the basis of positive-only evidence. However, due to the nonappearance of expletives in many cases we are left with a rather truncated hierarchy, excluding the subset principle as an effective learning procedure. Nonetheless, it is

still possible to consider that the values of the parameter are ordered in terms of markedness just as dictated by a subset hierarchy, as in theory the parameter is compatible with the subset principle. Equivalently, one can define the inclusion relations among values, rather than extensionally (i.e. among languages generated under these values), a possibility that follows naturally from the parameter. The markedness hierarchy and the learning procedure can then be defined accordingly (see Saleemi, in preparation, for definitions).

Recall that the set of null subject types under the four values progressively enlarges from value (a) to value (d): the set of null subjects under value (a) of the parameter is Ø; the set of possible null subjects under value (b) consists of nonarguments only; the set of possible null subjects under value (c) has as its members quasi-arguments as well as nonarguments; and the set of possible null subjects under value (d) contains nonarguments, quasi-arguments, and referential arguments. So, in this specific sense value (d) includes value (c), value (c) includes value (b), and value (b) includes value (a); that is, given the parameter P (= 5) with values ranging over P_a, ..., P_d, $P_a \subseteq P_b \subseteq P_c \subseteq P_d$.

The intuitive idea is that markedness is a function of certain internal properties of language, rather than of the external properties of particular languages (cf. Chomsky 1986a). However, it should be stressed that the data the child utilizes to instantiate parameters can only be described in terms of languages, an example of the intricate connection between parameter values and the corresponding languages. So let us now consider these languages and the set-theoretical relationships between them.

Let L(a) be the language generated by value (a) of the parameter; likewise L(b), L(c), and L(d). Then it is obvious that L(a) and L(b) intersect, as L(a) contains sentences with overt nonarguments, which L(b) does not have, and L(b) contains sentences with null nonarguments that are excluded by L(a). Next consider L(c). L(c) has sentences with null quasi-arguments, not included in L(b), whereas L(b) has sentences with overt quasi-arguments that are not contained in L(c). Therefore L(b) and L(c) also intersect. Now consider L(d). L(d) is coextensive with L(c) with respect to the nonpresence of nonreferential subjects, but it additionally contains referential null subjects. Consequently, $L(c) \subseteq L(d)$.

In short, the set-theoretical profile that emerges is rather mixed, incorporating *both* subset and intersecting relations, quite unlike what the subset principle and the related assumptions predict. The following consequence immediately ensues: if the correct language is any language other than L(a), then its selection might lead to overgeneralization *within* that language.

Suppose L(b) is the ambient language. Recall that in L(b) all argumental subjects must be overt. Then, when presented with

sentences with null nonarguments, the learner is bound to conjecture the parametric identity of the language. But notice that, given the no-negative evidence assumption, there is nothing that would prevent him from considering sentences with overt nonarguments to be in L(b), as exemplified here with respect to German.

(6) *Heute sind es zwei Kinder gekommen

Now consider that L(c) is the language to be learned. The presence in the data of sentences with both kinds of null nonreferential subjects should be sufficient to rule out L(a) and L(b), pointing to L(c) as the most likely choice. But the learner might still erroneously regard sentences with overt quasi-argument subjects (see the Yiddish example in 7), as well as those with overt nonargument subjects, to be in L(c).

(7) *Haynt geyt es a regn

Likewise in the case of L(d). The appearance in the data of null referential subjects should straightaway rule out L(a), L(b), and L(c). But the problem of possible overgeneralization to overt pleonastics within L(d) is still there. Considering that quite often expletives are homonymous with certain referential pronouns (e.g. Yiddish *es* and English *it* have referential analogues; notably, Welsh *hi* is 3rd person feminine singular), in principle overgeneralization can occur even though overt expletives are totally absent in the language being learned, as they are in Italian. In sum, it seems that although positive-only evidence is effective in *positively identifying* a language from among the four possible ones, it is not effective in *exactly identifying* that language.

One way out might be to posit that the child learner is able to undergeneralize within the conjectured language, by noticing the nonoccurrence of certain types of overt pleonastic subjects in the 'incomplete' data, in other words to resort to what Chomsky (1981) called indirect negative evidence (also see Lasnik, to appear, Oehrle 1985). Clearly, the use of indirect negative evidence can indeed exactly identify the correct value of a parameter from incomplete data.

It is possible to argue that indirect negative evidence is not really necessary, since alternative accounts based on positive-only evidence are available. For example, one can argue that the constraint stated in (8) is a part of the learner's *a priori* baggage, which stipulates the absence of redundant forms such as expletive subjects unless they are observed in positive data.

(8) *Redundancy Constraint*
Assume redundant forms to be absent unless they are exemplified in positive data.

This constraint, comparable but not identical to the Avoid Pronoun Principle, would make it possible for the parameter to be fixed from positive-only evidence. But there are a number of reasons why such a move might be ill-advised.

The constraint in question has no independent empirical status whatsoever, its only justification being that it salvages the 'no-negative evidence' condition; so it acts merely as a preemptive substitute for some form of (implicit) negative data. A major danger in stipulating such constraints is that it is all too easy to go on adding them to the learning system; the resulting interplay between parameters and constraints is very likely to become too undisciplined. Further, the constraint is essentially no more than a codified quasi-pragmatic tendency. I believe there are compelling methodological reasons for postulating that grammatical and pragmatic phenomena should never be conflated, since a maximally simple linguistic theory will be obtained by excluding these latter from the characterization of Universal Grammar.

Thus, it seems parsimonious to assume that the parameter is as stated in (5), and the extensional shortfall in the languages is due to pragmatic factors, not to be accounted for in the theory of grammar or learnability. The alternative to (8) (or similar *ad hoc* assumptions), of course, is to hypothesize that *some* use of indirect negative evidence is legitimate, so that the learner assumes pleonastics to be present unless they fail to show up in the data. Of course positive evidence must still retain its primary role, with indirect negative evidence supplementing this role only if the need to exclude some consequences of a parametric choice arises. So it is proposed that in general the following constraint holds of the human language learning system.

(9) *The Indirect Negative Evidence Constraint*
On no account can the choice of a parameter value, or a change in the value of a parameter, be made solely on the basis of indirect negative evidence.

4. Conclusion

To sum up, the relationship between parameter values and the languages they generate cannot invariably be captured extensionally in terms of simple set-theoretical relations, perhaps suggesting that parameters do not, strictly speaking, generate languages, but only fix the maximal bounds within which languages can be realized. Further, under the linguistic theory assumed the induction of a natural language is increasingly viewed as a search in a radically circumscribed choice space, rendering

it psychologically plausible that the child may be sensitive to nonoccurrence of certain types of data, a step which would be untenable under a less restrictive linguistic framework. Thus, in the end, the view of learnability just described may be empirically plausible only to the extent that the theory of principles and parameters is empirically plausible.

REFERENCES

Aoun, J., Hornstein, N., Lightfoot, D., and Weinberg, A. (1987). Two types of locality. *Linguistic Inquiry* 18.4:537-577.

Berwick, R. C. (1985). *The Acquisition of Syntactic Knowledge.* Cambridge, MA: MIT Press.

Bouchard, D. (1984). *On the Content of Empty Categories.* Dordrecht: Foris.

Chomsky, N. (1981). *Lectures on Government and Binding.* Dordrecht: Foris.

Chomsky, N. (1982). *Some Concepts and Consequences of the Theory of Government and Binding.* Cambridge, MA: MIT Press.

Chomsky, N. (1986a). *Knowledge of Language: Its Nature, Origin, and Use.* New York: Praeger.

Chomsky, N. (1986b). *Barriers.* Cambridge, MA: MIT Press.

Fabb, N. (1984). *Syntactic Affixation.* Doctoral dissertation, MIT.

Hyams, N. M. (1986). *Language Acquisition and the Theory of Parameters.* Dordrecht: Reidel.

Lasnik, H. (to appear). On certain substitutes for negative data. In *Essays on Restrictiveness and Learnability.* Dordrecht: Reidel.

Manzini, M. R. and Wexler, K. (1987). Parameters, binding theory, and learnability. *Linguistic Inquiry* 18.3:413-444.

Oehrle, R. T. (1985). Implicit negative evidence. Unpublished manuscript, Department of Linguistics, University of Arizona, Tucson.

Picallo, M. C. (1984). The Infl node and the null subject parameter. *Linguistic Inquiry* 15.1:75-102.

Rizzi, L. (1982). *Issues in Italian Syntax.* Dordrecht: Foris.

Rizzi, L. (1986). Null objects in Italian and the theory of *pro*. *Linguistic Inquiry* 17.3:501-557.

Safir, K. (1984). Missing subjects in German. In Toman, J. (ed). *Studies in German Grammar.* Dordrecht: Foris.

Safir, K. (1985). *Syntactic Chains.* Cambridge: CUP.

Saleemi, A. P. (in preparation). *Learnability and Parameter Fixation: The Problem of Learning in the Ontogeny of Grammar.* Doctoral dissertation, University of Essex.

Travis, L. (1984). *Parameters and Effects of Word Order Variation.* Doctoral dissertation, MIT.

Wexler, K. and Manzini, M. R. (1987). Parameters and learnability in binding theory. In Roeper, W. and Williams, E. (eds). *Parameter Setting.* Dordrecht: Reidel.